



China fertility survey 2017: design and implementation

Yaer Zhuang¹ · Shenghui Yang¹ · Jianan Qi¹ · Bohua Li¹ · Xueqing Zhao¹ · Zhili Wang^{1,2} 

Received: 27 September 2018 / Accepted: 28 November 2018 / Published online: 17 December 2018
© China Population and Development Research 2018

Abstract

This paper examines the design and the process used to carry out the China Fertility Survey 2017, a national representative survey that collected data on fertility desire, childbearing behavior, the use of childbearing services, and the determinants of childbearing behavior. The sampling method adopted was three-stage stratified probabilities proportional to size (PPS), and survey implementation made use of Computer Assisted Personal Interviewing (CAPI). CAPI played a significant role in survey design, last-stage sampling, interviewer training, face-to-face interviews, and questionnaire review and quality control. The survey results were compared with relevant data in the Integrated Management Information System for Population and Family Planning to check consistency. Ex post facto weighting was applied to correct sample structure bias. The process used to acquire accurate personal information is summarized. Suggestions based on consideration of sampling frame distortion by population mobility and other factors are put forward in the hope of improving similar sampling surveys in the future.

Keywords Fertility survey · Quality control · Ex post facto weighting

1 Introduction

China has conducted several national fertility and family planning surveys since the 1980s to understand childbearing behavior and contraceptive use among couples of reproductive age. These are the National Fertility Survey in 1982, the National Fertility and Contraception Survey in 1988, the National Family Planning Survey in 1992, the National Population and Reproductive Health Survey in 1997, the National Family Planning and Reproductive Health Survey in 2001, and the National Population and Family Planning Survey in 2006. In 2016, a new survey, the China Fertility

✉ Zhili Wang
zhiliw@126.com

¹ China Population and Development Research Center, Beijing, China

² Renmin University of China, Beijing, China

Survey, was proposed as a way to track changes that have occurred since birth policy reform began in 2013 and to assess needs for policy development to facilitate the birth policy reform. The survey took place in 2017 after approval by the National Bureau of Statistics of China.

The China Fertility Survey 2017 (hereafter CFS 2017) was conducted in a country that has experienced considerable demographic and social changes in the years since earlier fertility and family planning surveys were completed. There have also been developments in survey-related technology. Population mobility presents a big challenge to survey sampling and interviewing procedures. China had a migrant population of some 244 million people in 2017 (National Bureau of Statistics 2018), a historically unprecedented 17% of total population. Almost half of the migrants are women and most of these women are of reproductive age. As both urban and rural Chinese have become more aware of privacy concerns, conducting face-to-face interviews has become increasingly difficult. Also, procedures for reporting child births have been improved and there are more resources available to cross-check data, including the Population Information System and the Hospital Delivery Database, to name two. Today, computer techniques like CAPI and the internet are widely used in survey implementation. These provide more options for conducting interviews, assist with quality control, and improve survey data quality. In an effort to address emerging issues and new challenges while taking advantage of the latest developments in survey technology and effectively utilizing existing resources, the design and implementation of CFS 2017 were somewhat different than those of previous national surveys. As some scholars (Biemer and Lyber 2003) pointed out, a reliable process is the precondition for accurate data collection, and a satisfactory result can only be achieved by a satisfactory process. This paper will examine the design, implementation, and data checking and weight construction of CFS 2017 to provide information on the survey process and to assess the procedures and data quality.

2 Survey design

The main purposes of the China Fertility Survey 2017 were to provide evidence and references to support analysis of demographic changes and reforms that improve relevant family planning policies. As part of the effort to move forward reforms of family planning services and management, CFS 2017 offers insights into fertility changes in the last decade, the situation with respect to implementation of childbearing and childrearing public services in recent years, and the childbearing desire of couples. The survey sampling process and questionnaire were designed to serve these purposes.

2.1 Target population

The CFS 2017 was designed to sample current residents who were mainland Chinese females and aged 15–60 on July 1, 2017. The key definition is “current residents”,

it includes residents with local household registration and residents with household registration of other places while have been stayed 6 months or longer in current residency. Selecting survey respondents who are current residents has proven helpful in improving survey quality by significantly reducing the workload required to sample and implement face-to-face visits in comparison with other option which is selecting respondents from residents with local household registration only. Reproductive age is generally defined at 15–49 years, but since more than a decade had elapsed between the previous survey focused on fertility in 2006 and CFS 2017, the age range in the latter was extended to collect childbearing information from older women.

2.2 Questionnaire design

We designed two questionnaires for CFS 2017 data collection, an individual questionnaire and a community questionnaire. The individual questionnaire had four parts: childbearing intentions, childbearing behavior, the use of services related to childrearing, and determinants of childbearing. The community questionnaire was used to collect basic information about the community and services related to childbearing and childrearing.

A number of important questions in the individual questionnaire focus on childbearing behavior. Options of childbearing information collection include questions about whether the respondent has given birth in the last year (a question used by the population census in China), or detailed pregnancy history, or childbirth history. Studies have found that asking about pregnancy history is a good approach to eliminating the under reporting of live births (Marckwardt 1970). Given that one of the purposes of CFS 2017 was to review childbearing behavior during the 10 years period leading up to the survey, childbearing in the last year is obviously not an option. Childbirth history is simpler than pregnancy history and was used in previous surveys in China, such as National Fertility Survey in 1982. However, the use of pregnancy history is more appropriate from a reproductive health perspective, because information about abortion is included in the data. Some international surveys such as the World Fertility Survey and the Demographic and Health Survey collect pregnancy history information, and CFS 2017 is also designed to collect information on pregnancy history.

Information on childbearing desire and the use of services related to childbearing are also important parts of the questionnaire, and provide evidence that support efforts to improve relevant policies. Decisions and issues concerning childbearing are not simply a matter of the individual behavior of a woman, but are also influenced by the woman's family. As a result, the individual questionnaire gathered information about husbands and families to enable an in-depth study of childbearing desire and behavior.

To fully utilize CAPI techniques, we set ranges for responses when necessary and controlled for “skips” to ensure that responses were logical and reasonable. Additional CAPI functions, such as keeping the right sequence for pregnancy history

and remind respondents on the length of time between two pregnancies if it seems abnormally too short, are applied to eliminate under- or mis-reporting.

The questionnaire was tested several times at different survey sites, and problems found during this testing process were corrected.

2.3 Sampling design and procedures

To ensure that sample size was representative of the population nationally and regionally and was sufficient in size, after taking sampling errors into account, to estimate fertility, and given scheduling and budgetary constraints, CFS 2017 was designed for a sample of 250,000 women aged 15–60.

The sampling used a stratified three stage probabilities proportional to size (PPS) method. The survey covered all provinces and provincial level administrative units in mainland China. The sample was stratified by migrants and non-migrants under the jurisdiction of each provincial administrative unit, resulting in 64 strata. At each strata, the first stage of sampling was at the township level (the sampling unit was township, town, or sub-district, i.e. *xiang*, *zhen*, or *jiedao* in Chinese), the second stage was at the community level (the sampling unit was administrative village or neighborhood committee, i.e. *cunweihui* or *juweihui* in Chinese), and the third stage was at the individual level among the women aged 15–60.

The first and second stage sampling frames were produced from information reported by provincial administrations. In the first stage of sampling, sample size was pre-designed for each strata and implicitly stratified by administrative area and by township, town, and sub-district. The second stage of sampling was implicitly stratified by the residence situation (single dwelling or dormitory) and by the type of community (*cunweihui* or *juweihui*). The third stage of sampling was divided into two steps. The first step, carried out by the national implementation team, was to define the number of survey groups and randomly select survey groups based on the total number of eligible women in each community, and to distribute the list of survey groups by CAPI system to sample sites. The second step required local investigators to collect and report information on household identification number, age, and marital status of eligible women in the selected survey groups (100 women in each group), then in order of age and marital status, CAPI was used to perform a systematic sampling procedure that selected 20 women in each survey group. In this step, the investigators at each sampling site were trained to draw a grid map of dwelling units using internet map information in the CAPI system. Each grid referred to a “survey group”. Then these investigators were required to compile lists of households of random survey groups (selected by the national implementation team), and to confirm the accuracy of the lists with door-to-door visits and information on geographic location of the household was included with other data about the household.

If unanticipated problems made it impossible to use a survey site, a replacement survey site as similar as possible to the original was chosen at random. If a survey respondent had to be replaced for some reason, the CAPI system was used to select

a replacement of similar marital status and age. Of the original respondents selected for CFS 2017, 16,069 were replaced, a replacement rate is 6.4%.

2.4 Interviews

The CFS 2017 used two approaches to collecting information in the field: face-to-face interviews and online questionnaires. Although the response rate to online surveys is often lower than the response rate to face-to-face interviews, the online survey approach was appropriate in this case since the survey schedule overlapped with the school summer vacation period. As most women in school were unmarried and needed to provide less information for CFS 2017 than married women, asking them to fill out a questionnaire online after receiving a QR (quick response) code was acceptable.

Among the 246,840 face-to-face survey, 16,069 respondents had been replaced, and the total number of replaced respondents is 20,859 with a replacement rate of 6.43% and a nonresponse rate of 7.70%. The main reasons for the nonresponse were: (1) people had moved away (1.66%), (2) refusal (1.74%), (3) non-contact in all three visits (2.04%), ineligible (1.71%), break-off due to refusal (0.19%), break-off due to ineligible (0.24%) and veto (0.12%). We also developed a module of online survey specifically for the sample of 3160 students living in a centralized place. The survey was conducted with the assist of the head teachers or student counselors who randomly selected 20 students and provided them with the QR of online questionnaire. The students would fill in the questionnaire by themselves. Overall, 3106 students answered self-administered questionnaires and the nonresponse rate is 1.71%.

3 Survey implementation

3.1 Administrative management

Aware of the importance of the CFS 2017, government officials at the national, provincial and local levels worked to ensure the quality of the survey. The survey agenda was discussed in meetings of senior leaders from the National Health and Family Planning Commission of China (NHFPC), and several other meetings were devoted to identifying the purpose of the survey and to assigning management and implementation tasks for the survey. The leadership team, led by a vice minister from NHFPC, was composed of leaders from relevant departments and institutions, and was in charge of leadership and overall design of the survey. The office of the working group was in charge of coordination, implementation, monitoring, and management of the survey. A survey leadership group and 11 working groups were set up within the China Population and Development Research Center (CPDRC), the organization tasked with providing technical support for CFS 2017. These working groups were responsible for survey design, training, survey guidance and quality control, monitoring quality in the field, data checking, and data processing and analysis.

The relevant government departments were highly supportive of CFS 2017. As requested by the survey guidelines, each provincial Commission of Health and Family Planning set up a leadership group for the survey. Most provinces, autonomic regions and municipalities included the survey as one of the items in their annual performance evaluations for year 2017, and provided funding (which matched the special funds provided by the central government) for the survey. At the provincial level, one individual was in charge of coordinating and monitoring local work, as well as information distribution and communication. At the city and township levels, several supervisors were assigned to monitor the process and quality of local work. The supervisors were also responsible for monitoring survey data uploaded by investigators in the field, checking completed interviews and questionnaires, and identifying any feedback problems that cropped up.

A group of consultants comprised of experts from related fields such as population studies, statistics, and health was formed to provide advice at each stage to ensure that the survey was scientific and standardized. More than thirty consultancy meetings were organized to discuss issues such as questionnaire design (11 meetings), sampling design (9 meetings), weight construction (5 meetings), and preliminary results (5 meetings).

To standardize and regulate survey implementation, the “Technical Document of China Fertility Survey 2017” was developed as a guideline and reference for sampling frame development and sampling, questionnaire design, related indicator explanations and selections, and face-to-face interview procedures, as well as to support the training of investigators. “Plan for CFS 2017 Quality Control” and “Rules for Onsite Implementation of CFS 2017” were developed by CPDRC to control survey quality and to ensure that standardized procedures were used in the field. A consultation system was set up to deal with problems that emerged during survey process in a timely manner. Measures were taken to define and to recognize departments and individuals for good performance and to criticize poor performance.

3.2 Training

The CFS 2017 involved 12,500 interviewers and 3128 supervisors (in charge of guiding and monitoring onsite tasks including the interview process and quality). These people came from various health and family planning organizations or were local health and family planning workers. Interviewers had to be able to work with the internet and have a high school or higher education level. In consideration of questionnaire content and the target population of the survey, married women who met the recruitment criteria were given preference. It was felt married women would help to eliminate potential distance between interviewee and interviewer and make it easier for interviewees to answer sensitive questions. Among 12,500 interviewers selected, 85.7% were women, and the average age was 37. Most of the interviewers had previous survey experience and were quite capable of dealing with problems that sometimes arose during the interview process.

The training of interviewers, supervisors, and coordinators adapted a two-step approach. Step one was a national level effort to train the trainers, while step two

took place at the provincial level. CPDRC developed and produced training videos and curriculum, and carried out national level training. All interviewers, supervisors, and coordinators had to be trained at either the national or provincial level. CPDRC guided 16 provincial level training efforts, observed training processes and answered questions, and selected one site from each province to understand the third stage for sampling frame construction, setting up survey groups and maps of dwellings onsite, as well as construction of the name list of individuals in the survey group.

All training participants were required to take a test at the completion of training, and those who passed the test (with a score of 90 out of 100) were certified to perform a survey task. The CAPI system was used to generate a bank of test questions, including items about sampling content, questionnaire interviews, APP usage and quality control, and these were used to give the test scores. Among those who participated in the training and became supervisors or interviewers, 15.0% passed the test first time, 20.0% passed the test the second time they took it, and 65.0% passed the test the third time.

3.3 On line help

An online QQ group and a telephone hotline were set up by CPDRC to answer questions and to receive reports of problems that emerged during the survey process. All problems were reviewed and summarized, and possible methods of dealing with the problems were written up as technical support documents for distribution nationwide. QQ groups were set up at each level to facilitate discussions of open questions, problem solving and other communications, and to share experiences as a way to prevent problems from arising.

3.4 Field work

CFS 2017 included 12,500 survey sites (communities) from 6078 townships, towns or sub-districts under the jurisdiction of 2737 counties, cities or districts. The interviewers worked very hard to complete their tasks within deadline; this was especially challenging for those who worked in remote rural areas. Interviews often took place after work hours with 38.2% of the questionnaires or 94,356 of the 246,840 face-to-face interviews completed during the evenings or weekends. Communications with interviewees and scheduling interview times in advance resulted in a high response rate; the non-response rate was only 1.9%.

The CAPI system played a significant role in quality control. It provided functions that allowed for making voice recordings of interviews and verifying interview locations. This prevented the skipping of face-to-face interviews, having proxies fill out questionnaire, or other kinds of malfeasance. CAPI was also used to perform an almost “simultaneous” data checks immediately after the completed questionnaire was uploaded. Problems were identified quickly, and feedback and corrections could be obtained from the interviewer in a timely manner. This process helped to prevent the occurrence of similar problems in the future.

3.5 Monitoring and quality control

Monitoring and supervising the survey process is very important to quality control. A standardized survey is less biased if interviewers strictly follow guidelines, and effective monitoring and supervision can further eliminate errors (Hao 2004, p. 196). NCHFP organized several expert monitoring and supervisory groups. The groups visited survey sites in 16 provinces, autonomous regions and municipalities to understand how teams of supervisors and interviewers were structured, and to learn about how interviews were approached and about the attitude of respondents. The groups monitored the progress and quality of survey implementation, and identified problems that emerged during survey implementation. The monitoring and supervisory groups observed interviews, revisited interviewees, organized focus group discussions, discussed difficulties and dealt with problems. They reported on the situation in the field, and provided feedback on problems in an effort to prevent problems from reoccurring.

4 Data comparison and the construction of weights

Post hoc evaluation is an approach widely used to assess the quality of survey data. Post hoc quality checks and data comparisons with other sources are commonly used in evaluations. Both methods were applied to CFS 2017 to check the quality of survey data. Post hoc quality checks require a carefully designed work plan and strong investigator teams to collect reliable reference data, but they do not need a large sample, usually 5–10% of respondents are revisited (for example, as in the National Population Census, see: Wu 2002). Data comparisons with data from other sources can be comparisons of whole samples or of sub-samples.

4.1 Post hoc evaluation

For the post hoc evaluation, 64 survey sites out of 12,500 nationwide (5.12%) were selected for re-visit interviews, and 2833 respondents were selected for re-interview by telephone. Re-visit interviews focused on basic information (birth date, marital status, date of first marriage, household type, education), childbearing information (total number of children born, number of sons and number of daughters, birth dates, sex, and survival status of each child).

The result of re-visit face to face interviews showed that home visits, the replacement of respondents, and the monitoring and supervision of the survey process were implemented according to requirements. The consistency rates for birth dates, marital status, and total number of children born were 96.6%, 98.0%, and 96.3%, respectively. The result of re-interviews by telephone showed a high consistency rate of 98.9% for the total number of children born.

4.2 Data comparison

The individual information collected for CFS 2017 was compared with data sets from Integrated Management Information System for Population and Family Planning. The identification numbers of individuals were used for matching, and when cases of inconsistent information were identified, the survey data was returned to the field site for re-checking and re-confirmation. The rate of successful matching was 88.4% (220,828 women out of 249,946 women). The cause of miss-matches was mainly due to no report being available or the miss-reporting of identification numbers.

Results from data comparisons of the total number of children born were divided into three groups: the total number of children born found by the survey was either less, equal to, or more than the number in the comparison database. Among cases with identification matching with database, the proportion of “equal” cases was 72.6%, proportion of “less” cases was 3.6%, and proportion of “more” cases was 23.8%. After checking the 8011 “less” cases at survey sites, 3140 were identified as under-reporting of births (which represented 0.8% of total children born in the CFS 2017); the “more” cases included 280 cases of over-reporting. The principal reasons for under-reporting were deliberate under-reporting, the child died, or the child went to live with the father after a divorce. Over-reporting was often due to a man remarriage and bringing children from a previous marriage or adopted children to the new marriage. Some “more” cases were the result of databanks not having been updated at the time of the survey (July 1, 2017).

4.3 Ex post facto weighting

We constructed sampling design weights to ensure to the greatest extent possible that the sample population was representative of the whole population covered. At the individual level, the weighting helped to adjust for non-responses. However, population migration and other factors created challenges affecting the completeness of sampling frames mainly at the third stage. Furthermore, because the construction of sampling frames adopted a top-down approach that was highly dependent on existing local statistics (that were sometimes not up to date), it was possible for the size of a sampling unit to vary after the unit was selected. Most importantly, in the third stage of sampling frame construction, the reported target population was likely to be group with more older women and fewer younger women who had never married, since young, unmarried women are highly mobile and difficult to enumerate. For example, after weighting for sampling design and non-responses, the proportions of women aged 15–34, of women who have never married, and of women from non-agricultural households in the sample are, not surprisingly, lower than the proportion of these groups of women in the overall population. Because these features are important in estimating fertility, ex post factor weighting is necessary. This weighting process included in the survey design incorporated lessons learned from sampling frame bias in previous surveys.

The weights we constructed for age structure referred to household registration information provided by the Ministry of Public Security. The iteration method was

used to determine marital status of the sample, based on provincial data from the 2015 National 1% Population Sampling Survey. We constructed weights for *hukou* by estimating the proportion of non-agriculture women aged 15–60 from the years 2006 to 2017, based on data from the 2000 and 2010 population censuses. The proportions of women aged 15–34, women who had never married, and women from non-agriculture households were 40.9%, 20.7%, and 33.2%, respectively, after weighting. These proportions were consistent with the relevant proportions in the overall population.

5 Conclusion and discussion

This report reviews the entire process used for the China Fertility Survey 2017, including survey design, implementation, data processing and quality control. We conclude that the survey results were accurate and reliable, while bias caused by sample structure was corrected after ex post facto weighting. Effective management and strong leadership, standardized operational guidelines, training, onsite monitoring and technical support, ex post quality checks, and data comparisons with other sources all helped to ensure that the survey was successfully implemented and the results of high quality. The use of CAPI played a very important role in quality control throughout all procedures, and improved the effectiveness of data collection and the quality of data collected.

Data quality was dependent on the reliability of responses and having an unbiased sample structure. Interviewers who act responsibly are important in such a large scale national survey; however, process monitoring of, for example, geographic allocations during sampling frame construction and face-to-face interviews, is also necessary for quality control. Population mobility and other factors create enormous challenges to the task of constructing a complete and accurate sampling frame. Some of the sampling frame construction process in this survey did not exactly follow the guidelines, and this was a major cause of sampling frame defects in the survey.

The problems identified in sampling frame construction for this survey could be encountered in similar surveys in the future. We suggest improvements and adjustments to the construction method for sampling frames for fertility sampling surveys. First, related population parameters such as age structure, marital status structure, and the proportion of migrants should be included in the sampling plan to check for possible bias of the sampling structure, and to provide references for ex post factor weighting. Second, because marital structure is very important to estimates of fertility estimation, stratification by marital status could be an option during the final stage. Third, because a majority of young, never married women live in dormitories, especially those provided by schools, information on students in high school, college or university could be collected from education departments. Fourth, we suggest performing quick summaries of sampling frame data and sample data at each stage to assess the quality of the sampling frames and onsite sampling. This kind of activity would facilitate identification of problems that could be corrected relatively early in the process.

It is worth noting that the data was weighted at the national level according to women's age, marriage status, and their household registration status (the proportion of non-agricultural hukou). However, results at provincial level could not be estimated by weighting only. Moreover, some fertility indicators should be computed with enough sampling size and less relative errors. Although the data could be applied to compute the national-level indicators, it is incapable for the estimation of provincial-level indicators, especially for complicated indicator such as TFR. Therefore, when the data is used to compute provincial-level indicators and below, the weights should be adjusted with more relevant information. Only when the weights are calibrated and relative errors are controlled by larger sample size, the data could be applied regionally and locally.

References

- Biemer, P., & Lyber, L. (2003). *Introduction to Survey Quality*. Hoboken: Wiley.
- Hao, D. (2004). *Social Survey Method*. Beijing: Renmin University Press.
- Marckwardt, A. (1970). Descripción y resultados preliminares de la encuesta nacional de fecundidad. Paper presented at the Second Symposium on Household Surveys in Latin America, Rio de Janeiro, Brazil.
- National Bureau of Statistics of China. (2018). 2017 National Statistical Report on Economic and Social Development. Feb. 28. http://www.stats.gov.cn/tjsj/zxfb/201802/t20180228_1585631.html, Accessed at 20 Aug 2018.
- Wu, J. (2002). The Post Enumeration Survey in the Census. *South China Population*, 3, 18–24.



Yaer Zhuang is research fellow of China Population and Development Research Center, Beijing, China. She is the first author of this article.



Shenghui Yang is associate research fellow of China Population and Development Research Center, Beijing, China.



Jianan Qi is associate research fellow of China Population and Development Research Center, Beijing, China.



Bohua Li is research fellow of China Population and Development Research Center, Beijing, China.



Xueqing Zhao is research assistant of China Population and Development Research Center, Beijing, China.



Zhili Wang is associate research fellow of China Population and Development Research Center, and is a doctoral student of Renmin University of China, Beijing, China. He is the correspondent author of this article.